Individual based and mean-field modelling of direct aggregation

Martin Burger[1]   Jan Haškovec[2]   Marie-Therese Wolfram[3]

Abstract. We introduce two models of biological aggregation, based on randomly moving particles with individual stochasticity depending on the perceived average population density in their neighbourhood. In the first-order model the location of each individual is subject to a density-dependent random walk, while in the second-order model the density-dependent random walk acts on the velocity variable, together with a density-dependent damping term. The main novelty of our models is that we do not assume any explicit aggregative force acting on the individuals; instead, aggregation is obtained exclusively by reducing the individual stochasticity in response to higher perceived density. We formally derive the corresponding mean-field limits, leading to nonlocal degenerate diffusions. Then, we carry out the mathematical analysis of the first-order model, in particular, we prove the existence of weak solutions and show that it allows for measure-valued steady states. We also perform linear stability analysis and identify conditions for pattern formation. Moreover, we discuss the role of the nonlocality for well-posedness of the first-order model. Finally, we present results of numerical simulations for both the first- and second-order model on the individual-based and continuum levels of description.

Key words: Direct aggregation, Density dependent random walk, Degenerate parabolic equation, Mean field limit.

arXiv:1112.1055v2 [math.DS] 21 Feb 2012

---

[1]Institut für Numerische und Angewandte Mathematik, Westfälische Wilhelms-Universität Münster, Einsteinstr. 62, 48149 Münster, Germany; e-mail: *martin.burger@wwu.de*

[2] King Abdullah University of Science and Technology, Thuwal, KSA; e-mail: *jan.haskovec@kaust.edu.sa*

[3]Faculty of Mathematics, Universität Wien, Nordbergstrasse 15, A-1090 Wien, Austria and DAMTP, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA, Great Britain; e-mail: *marie-therese.wolfram@univie.ac.at*

1 Introduction

Animal aggregation is the process of finding a higher density of animals at some place compared to the overall mean density. Its formation may be triggered by some environmental heterogeneity that is attractive to animals (the aggregate forms around the environmental template), by physical currents that trap the organisms through turbulent phenomena, or by social interaction between animals [7, 19]. Aggregation may serve diverse functions such as reproduction, formation of local microclimates, anti-predator behaviour (see for instance [32] for a study of reducing the risk of predation to an individual by aggregation in *Aphis varians*), collective foraging and much more (see, e.g., [14] for a relatively recent survey). Aggregation also plays an important role as an evolutionary step towards social organization and collective behaviours [15]. These aspects explain the continuing interest in understanding not only the function of animal aggregation, but also the underlying mechanisms. The development of quantitative mathematical models is an essential part in this quest. While such models first of all help to link individual behavioural mechanisms to spatio-temporal group patterns, they also often aim at explaining the observed dynamic efficiency of animal groups to adapt to environmental variation, and, moreover, provide a valuable tool to study in more detail the system dynamics and their robustness to variations in individual behavioural parameters or external parameters, see, e.g., [11].

Going back to the pioneering work of Skellam [29], continuum spatiotemporal population dynamics are traditionally modelled by reaction-convection-diffusion PDEs and systems thereof. In these models, diffusion typically describes the avoidance of crowded areas by the individuals, and as such acts as an anti-aggregative force, working against the typically aggregative effect of convection, see the surveys [22] and [21]. Our work goes into the opposite direction and shows that biological aggregations can be a consequence of solely random, diffusive motion of individuals, who respond to the local population density observed in their neighborhood by increasing or decreasing the amplitude of their random motion. This kind of behaviour was observed in insects, for instance the pre-social German cockroach Blattella germanica. These animals are known to be attracted to dark, warm and humid places [26]. However, the works by Jeanson et al. [9, 10] have shown that cockroach larvae also aggregate in the absence of any environmental template or heterogeneity. In this case aggregation is the result of social interactions and happens as a self-organized process. The individual based model developed and parameterized in [10] identified a simple decisive mechanism that can be summarised in the following way: cockroaches do not rest for a long time in places with few conspecifics, and once moving they stop preferentially in places of high cockroach density. A mathematically better tractable version of this decisive mechanism is the model of individuals performing random walks with density dependent coefficients. In the continuum limit, this leads to degenerate diffusion models, which, however, depending on their parameters and data, might have ill-posed regimes. In particular, Turchin [31] derived a 1D model for a population with density $u(x,t)$ of the form

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\left[\phi'(u)\frac{\partial u}{\partial x}\right], \tag{1.1}$$

where $\phi(u) = (\mu/2)u - k_0 u^2 + (2k_0/2\omega)u^3$ with the positive coefficients μ, k_0 and ω. Turchin used the above equation to describe the aggregative movement of *Aphis varians*, a herbivore of fireweeds (*Epilobium angustifolium*). He also discussed the possible ill-posedness of some initial and boundary value problems associated with (1.1). Depending on the actual profile of ϕ, Turchin also classified two different types of aggregation. Essentially the same model has been derived independently in [1] as a model of cell motility with volume filling and cell-to-cell adhesion. The authors showed that the diffusivity can become negative

2

if the cell adhesion coefficient is sufficiently large and related this to the presence of spatial oscillations and development of plateaus in their numerical solutions of the discrete model. Moreover, they used a combination of stability analysis of the discrete equations and steady-state analysis of the limiting PDE to gain better understanding of the qualitative predictions of the model. Another work studying an equation of the type (1.2) is [23], where existence and uniqueness of solutions of the initial value problem with homogeneous Neumann boundary conditions in a bounded domain of $\mathbb{R}^n$ was shown and some aspects of the aggregating behaviour were studied analytically.

In [25], following [18], the authors provide an alternative derivation of (1.1) for low population density, using a biased random walk approach. Then they study the existence of traveling wave solutions for a purely negative or zero diffusion equation with a logistic rate of growth $g(u)$,

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\left[\phi'(u)\frac{\partial u}{\partial x}\right] + g(u)\,, \tag{1.2}$$

discuss the well- and ill-posedness of certain boundary conditions associated with some purely negative diffusion equations with logistic-like kinetic part and provide some numerical examples.

One possibility to overcome the difficulty of the possible ill-posedness of (1.1) and (1.2) is to introduce a nonlocality, i.e., to substitute the term $\Delta\phi(u)$ by ΔJ with

$$J(x,t) = \int_\Omega K(x,y)\phi(u(y,t))\,\mathrm{d}y$$

with a nonnegative kernel $K(x,y)$. A particular choice made in [6] is to define $K(x,y)$ as the Green function of $(I - \lambda\Delta)$ for a constant $\lambda > 0$. Equation (1.2) becomes then

$$\frac{\partial u}{\partial t} = \Delta[I - \lambda\Delta]^{-1}\phi(u) + g(u)\,,$$

which is equivalent to

$$\frac{\partial u}{\partial t} = \Delta\left(\phi(u) - \lambda g(u) + \lambda\frac{\partial u}{\partial t}\right) + g(u)\,.$$

This equation was studied in [24] and can be considered as a model of aggregating population with a migration rate determined by ϕ and total birth and mortality rates characterized by g. In [24] it was shown that the aggregating mechanism induced by $\phi(u)$ allows for survival of a species in danger of extinction and performed numerical simulations suggesting that, for a particular version of $\phi(u)$, the solutions stabilize asymptotically in time to a not necessarily homogeneous stationary solution. Other two works going in this direction are [30], which we discus later (Section 4.4), and the model of home range formation in wolves due to scent marking [2],

$$\begin{aligned}\frac{\partial u}{\partial t} &= \Delta\left(\frac{D_0 u}{1 + p/\alpha}\right)\,,\\ \frac{\partial p}{\partial t} &= u(\gamma + m(p)) - \mu p\,.\end{aligned}$$

Here $u(x,t)$ is the population density of wolves, $p(x,t)$ the density of their scent marks and D_0, γ, α and μ positive parameters. The increasing function $m(p)$ describes enhanced scent mark rates in the presence of existing scent marks. A nonlocal version of the model with m depending on the averaged version of

p is also considered. The authors of [2] show that the model produces distinct home ranges; in this case the pattern formation results from the positive feedback interaction between the decreased diffusivity of wolves in the presence of high scent mark densities and the increased production of new scent marks in these locations.

In our paper we introduce and study two models where formation of aggregates results from random fluctuations in the population density and is supported merely by reducing the amplitude of the individual random motion in response to higher perceived density. This leads to nonlocal individual-based and PDE models, where the nonlocality stems from calculating the perceived density as a weighted average over a finite sampling radius. This is usual in modelling biological interactions, see for instance [12, 7, 30, 20] and many more. Our models were inspired by the above mentioned observations of German cockroach [9, 10], but does not aim to be a realistic description of their social behaviour. Instead, we consider our work as a proof of concept, where the main characteristic of our models is that we do not assume any deterministic interaction between the individuals that would actively push them to aggregate. This approach was followed for instance in stochastic run-and-tumble models of chemotaxis [27]. Closely related to our work is the series of papers [8, 16, 17]. However, only the first-order model under specific conditions was studied there. The new aspects contributed by our work are the generality of the models (first- and second-order), more rigorous derivation of the mean-field limits based on the generalized BBGKY-hiearchy, and rigorous well-posedness and stability analysis of the first-order model.

Our paper is structured as follows: In Section 2 we introduce two stochastic individual based models, where every individual is able to sense the average population density in its neighborhood, and respond in terms of increased or decreased stochasticity of its movement. In particular, the diffusivity is reduced in response to higher perceived density. We present a first-order model, where the location of each individual is subject to a density-dependent random walk, and a second-order model, where the density-dependent random walk acts on the velocity variable, together with a density-dependent damping term. The advantage of the second-order model is that it is possible to introduce a cone of vision, which depends on the direction of each individual's movement. In Section 3 we formally derive the corresponding mean-field limits, leading to a nonlocal, nonlinear diffusion equation for the first-order model, and a nonlinear Fokker-Planck kinetic equation for the second-order model. Moreover, we show that the diffusive limit of the kinetic equation leads to the first-order diffusion equation derived before. Then, in Section 4, we show the existence of weak and measure-valued solutions of the first order mean-field model and perform linear stability analysis of the uniform steady states, finding regimes that correspond to the sought-for pattern formation (direct aggregation). Finally, in Section 5 we present the results of numerical simulations of our models. We performed Langrangian simulations of the first- and second-order agent-based models in a periodic 2D domain, Eulerian simulations of the first-order mean-field model in 1D and 2D periodic domains, and, finally, of the second-order mean-field model in a spatially 1D periodic domain with 1D velocity.

2   The stochastic individual based models

We consider the set of $N \in \mathbb{N}$ individuals with time-dependent positions $x_i(t) \in \mathbb{R}^d$, $i = 1, \ldots, N$, with $d \geq 1$. Every individual is able to sense the average density of other individuals in its neighborhood, given by

$$\vartheta_i(t) = \frac{1}{N} \sum_{j \neq i} W(x_i - x_j), \tag{2.1}$$

4

where $W(x) = w(|x|)$ with $w : \mathbb{R}^+ \to \mathbb{R}^+$ is a bounded, nonnegative and nonincreasing weight, integrable on $\mathbb{R}^d$. A generic example of w is the characteristic function of the interval $[0, R]$, corresponding to the sampling radius $R > 0$. The average density ϑ_i is then simply the fraction of individuals located within the distance R from the i-th individual:

$$\vartheta_i(t) = \frac{1}{N} \#\{j; |x_i - x_j| \leq R\}.$$

For the passage to the mean-field limit $N \to \infty$ and the analysis of the resulting equations, we will have to pose certain smoothnes assumptions on W, which actually exclude the choice of w to be a characteristic function. However, from the modelling point of view this is not a concern.

Let us now introduce our two models:

- In the first-order model the individual locations are subject to average density-dependent random walk,

$$\mathrm{d}x_i = G(\vartheta_i)\,\mathrm{d}B_i^t, \qquad i = 1, \ldots, N, \tag{2.2}$$

  where B_i^t are independent d-dimensional Brownian motions and $G : \mathbb{R}^+ \to \mathbb{R}^+$ is a bounded, nonnegative and nonincreasing function. For instance, in the numerical simulations of Section 5 we use $G(s) = \exp(-s/3)$. However, we as well allow for degeneracy, where $G(s) = 0$ for all $s \geq s_0 > 0$.

- In the second-order model the individuals are described not only by their locations $x_i(t) \in \mathbb{R}^d$, but also by the velocities $v_i(t) \in \mathbb{R}^d$, $i = 1, \ldots, N$. The advantage of this description, in contrast to the first-order model, is that every individual has a well defined direction of movement. Therefore, the weight W in the calculation of the average densities ϑ_i can also depend on the relative angle with respect to the individual's direction of movement,

$$\vartheta_i(t) = \frac{1}{N} \sum_{j \neq i} W(x_i - x_j, v_i), \tag{2.3}$$

with $W(x, v) = w\left(|x|, \frac{x \cdot v}{|x||v|}\right)$. This is important from the modeling point of view, since we can define not only the sampling radius $R > 0$, but also a restricted cone of vision with angle $\alpha \in (0, \pi]$; we set then $w(s, z) = \chi_{[0,R]}(s)\chi_{[\cos \alpha, 1]}(z)$.

The velocity in our model is subject to a density-dependent random walk and a density-dependent damping term:

$$\mathrm{d}x_i \;=\; v_i\,\mathrm{d}t, \tag{2.4}$$
$$\mathrm{d}v_i \;=\; -H(\vartheta_i)v_i\,\mathrm{d}t + G(\vartheta_i)\,\mathrm{d}B_i^t. \tag{2.5}$$

The function G is as before, while $H : \mathbb{R}^+ \to \mathbb{R}^+$ is a nonnegative and nondecreasing function. The damping term $-H(\vartheta_i)v_i$ is introduced in order to slow down the agents' motion when they approach a crowded place. Obviously, this mechanism is not only necessary to allow for formation of aggregates, but also very natural from the modeling point of view.

3 The mean-field limits

3.1 The first-order model

We start with the derivation of the mean-field limit of the first-order model (2.2). Unfortunately, the standard framework of BBGKY hierarchies cannot be applied, since due to the structure of the model it is impossible to obtain a hierarchy where the evolution of the k-th marginal is expressed in terms of a finite number of higher-order marginals. Therefore, we have to apply the recently developed technique of introduction of auxiliary variables [3] and their elimination after the mean-field limit passage.

In particular, under the additional assumption $W \in C^1(\mathbb{R}^d)$, we extend the model by introducing the average densities ϑ_i given by (2.1) as a new set of independent variables, governed by the system of stochastic differential equations

$$\mathrm{d}\vartheta_i = \frac{1}{N} \sum_{j \neq i} \nabla W(x_i - x_j) \left(G(\vartheta_i) \, \mathrm{d}B_i^t - G(\vartheta_j) \, \mathrm{d}B_j^t \right) . \tag{3.1}$$

Let us point out that the random walks B_i^t, B_j^t in (3.1) are correlated with those in (2.2). Using the Itô formula, we turn to the equivalent formulation of the stochastic system (2.2), (3.1) in terms of the corresponding Fokker-Planck equation

$$
\begin{aligned}
\frac{\partial f^N}{\partial t} &= \frac{1}{2} \sum_{i=1}^{N} \Delta_{x_i} \left(G(\vartheta_i)^2 f^N \right) + \sum_{i=1}^{N} \frac{\partial}{\partial \vartheta_i} \nabla_{x_i} \cdot \left(\frac{1}{N} \sum_{j \neq i} \nabla W(x_i - x_j) G(\vartheta_i)^2 f^N \right) \\
&\quad - \frac{1}{N} \sum_{i=1}^{N} \sum_{j \neq i} \sum_{k \neq i} \frac{\partial}{\partial \vartheta_j} \nabla_{x_i} \cdot \left(\nabla W(x_k - x_i) G(\vartheta_i)^2 f^N \right) \\
&\quad + \frac{1}{2N^2} \sum_{i=1}^{N} \sum_{j \neq i} \sum_{k \neq i} \frac{\partial^2}{\partial \vartheta_i^2} \left(\nabla W(x_i - x_k) \cdot \nabla W(x_i - x_l) G(\vartheta_i)^2 f^N \right) \\
&\quad + \frac{1}{2N^2} \sum_{i=1}^{N} \sum_{j \neq i} \frac{\partial^2}{\partial \vartheta_i^2} \left(|\nabla W(x_i - x_j)|^2 G(\vartheta_j)^2 f^N \right) \\
&\quad + \frac{1}{2N^2} \sum_{i=1}^{N} \sum_{j \neq i} \sum_{k \neq i,j} \frac{\partial^2}{\partial \vartheta_i \partial \vartheta_j} \left(\nabla W(x_i - x_k) \cdot \nabla W(x_j - x_k) G(\vartheta_k)^2 f^N \right) \\
&\quad - \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j \neq i} \sum_{k \neq i,j} \frac{\partial^2}{\partial \vartheta_i \partial \vartheta_j} \left(\nabla W(x_i - x_k) \cdot \nabla W(x_i - x_j) G(\vartheta_i)^2 f^N \right) ,
\end{aligned}
$$

where $f^N = f^N(t, x_1, \vartheta_1, \ldots, x_N, \vartheta_N)$ is the N-particle distribution function. Defining the k-particle marginals f_k^N by

$$f_k^N(t, x_1, \vartheta_1, \ldots, x_k, \vartheta_k) = \int_{\mathbb{R}^{d(N-k)}} \int_{(0,\infty)^{N-k}} f^N(t, x_1, \vartheta_1, \ldots, x_N, \vartheta_N) \, \mathrm{d}\vartheta_{k+1} \ldots \vartheta_N \, \mathrm{d}x_{k+1} \ldots \mathrm{d}x_N ,$$

we obtain the so-called BBGKY-hierarchy for the system $(f_k^N)_{k=1}^N$. In particular, the equation for $f_1 =$

$f_1(t, x, \vartheta)$ reads

$$
\begin{aligned}
\frac{\partial f_1^N}{\partial t} &= \frac{1}{2}\Delta_x \left(G(\vartheta)^2 f_1^N \right) + \frac{\partial}{\partial \vartheta}\nabla_x \cdot \left(G(\vartheta)^2 \int_{\mathbb{R}^d} \int_0^\infty \nabla W(x-y) f_2^N(x, \vartheta, y, \sigma)\, \mathrm{d}\sigma\, \mathrm{d}y \right) \\
&+ \frac{1}{2}\frac{\partial^2}{\partial \vartheta^2} \left(G(\vartheta)^2 \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \int_0^\infty \int_0^\infty \nabla W(x-y)\nabla W(x-z) f_3^N(x, \vartheta, y, \sigma, z, \tau)\, \mathrm{d}\sigma\, \mathrm{d}\tau\, \mathrm{d}y\, \mathrm{d}z \right) \\
&+ \frac{1}{2N}\frac{\partial^2}{\partial \vartheta^2} \left(\int_{\mathbb{R}^d} \int_0^\infty G(\vartheta)^2 |\nabla W(x-y)|^2 f_2^N(x, \vartheta, y, \sigma)\, \mathrm{d}\sigma\, \mathrm{d}y \right),
\end{aligned}
\tag{3.2}
$$

where we, for the sake of legibility, dropped the indices at x_1 and ϑ_1. Now, we pass formally to the limit $N \to \infty$, assuming that $\lim_{N\to\infty} f_k^N = f_k$ for all $k \geq 1$. Moreover, we admit the usual *molecular chaos* assumption about vanishing particle correlations as $N \to \infty$,

$$
f_k(t, x_1, \vartheta_1, \ldots, x_k, \vartheta_k) = \prod_{i=1}^k f_1(t, x_i, \vartheta_i) \qquad \text{for all } k \geq 2 .
$$

Then, one obtains from (3.2) the one-particle equation

$$
\frac{\partial f_1}{\partial t} = \frac{1}{2}\Delta_x \left(G(\vartheta)^2 f_1 \right) + \frac{\partial}{\partial \vartheta}\nabla \cdot \left(G(\vartheta)^2 (\nabla W \circledast f_1) f_1 \right) + \frac{1}{2}\frac{\partial^2}{\partial \vartheta^2} \left(G(\vartheta)^2 |\nabla W \circledast f_1|^2 f_1 \right),
\tag{3.3}
$$

where the operator $\circledast$ is defined as

$$
\nabla W \circledast f_1(t, x) = \int_{\mathbb{R}^d} \int_0^\infty \nabla W(x-y) f_1(t, y, \vartheta)\, \mathrm{d}\vartheta\, \mathrm{d}y .
$$

Finally, we reduce (3.3) to obtain the standard mean-field description of the system (2.2) by removing the auxiliary variable ϑ. Indeed, a relatively lengthy, but straight-forward formal calculation shows that (3.3) posesses weak solutions of the form

$$
f_1(t, x, \vartheta) = \varrho(t, x)\delta(\vartheta - W * \varrho(t, x)), \qquad \text{with } W * \varrho(t, x) = \int_{\mathbb{R}^d} W(x-y)\varrho(t, y)\, \mathrm{d}y
$$

and $\varrho = \varrho(t, x)$ satisfies the nonlinear diffusion equation

$$
\frac{\partial \varrho}{\partial t} = \frac{1}{2}\Delta_x \left(G(W * \varrho)^2 \varrho \right).
\tag{3.4}
$$

3.2 The second-order model

With the same procedure as before we derive the formal mean-field limit of the model (2.4)–(2.5). We omit the details here and immediately give the resulting kinetic Fokker-Planck equation for the particle distribution function $f = f(t, x, v)$,

$$
\frac{\partial f}{\partial t} + v \cdot \nabla_x f = \nabla_v \cdot \left(H(W \circledast f)vf + \frac{1}{2}\nabla_v \left(G(W \circledast f)^2 f \right) \right),
\tag{3.5}
$$

where, with a slight abuse of notation, the convolution operator $\circledast$ is defined as

$$
W \circledast f(t, x, v) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} W(x-y, v) f(t, y, w)\, \mathrm{d}w\, \mathrm{d}y .
$$

Let us make the following observation: If the weight W does not depend on v, such that $W \circledast f$ does not depend on v as well and $W \circledast f(t, x) = W * \varrho(t, x)$, we can write the (non-closed) system for the mass, momentum and energy densities

$$\varrho(t, x) = \int_{\mathbb{R}^d} f(t, x, v) \, \mathrm{d}v \,, \qquad \varrho u(t, x) = \int_{\mathbb{R}^d} f(t, x, v) v \, \mathrm{d}v \,, \qquad \varrho E(t, x) = \frac{1}{2} \int_{\mathbb{R}^d} f(t, x, v) |v|^2 \, \mathrm{d}v \,, \quad (3.6)$$

associated with the solution f of (3.5) as

$$\frac{\partial \varrho}{\partial t} + \nabla_x \cdot (\varrho u) = 0 \,, \tag{3.7}$$

$$\frac{\partial (\varrho u)}{\partial t} + \nabla_x \cdot \left(\int_{\mathbb{R}^d} f(t, x, v) v \otimes v \, \mathrm{d}v \right) = -H(W * \varrho) \varrho u \,, \tag{3.8}$$

$$\frac{\partial (\varrho E)}{\partial t} + \nabla_x \cdot \left(\frac{1}{2} \int_{\mathbb{R}^d} f(t, x, v) |v|^2 v \, \mathrm{d}v \right) = -2H(W * \varrho) \varrho E + \frac{d}{2} G(W * \varrho)^2 \varrho \,. \tag{3.9}$$

We observe that only mass is conserved, while momentum and energy are not (neither locally nor globally). Indeed, the momentum is dissipated due to the "friction" term, whose strength depends non-locally on ϱ due to $H(W * \varrho)$. The energy is, on one hand, dissipated due to the same term, on the other hand is created due to the diffusive term in (3.5), at the rate $\frac{d}{2E} G(W * \varrho)^2$. In equilibrium, we have $H(W * \varrho) \varrho u = 0$, which means that we either have empty regions ($\varrho = 0$) or regions with positive density, but zero velocity. In these populated regions the equilibrium energy is given by

$$E = E[\varrho] = \frac{d}{4} \frac{G(W * \varrho)^2}{H(W * \varrho)} \,.$$

3.3 Diffusive limit of the second-order model (3.5)

We show that the first-order model (3.4) is obtained from (3.5) in the formal diffusive limit, under the assumption that the weight W does not depend on v, which we make for all of this section. Let us recall that then $W \circledast f$ does not depend on v as well and $W \circledast f(t, x) = W * \varrho(t, x)$, with ϱ defined by (3.6).

We start by observing that the equilibria of the collision operator of (3.5) are given by the local Maxwellians

$$\mathcal{M}[\varrho](v) = \left(\frac{H(W * \varrho)}{2\pi G(W * \varrho)^2} \right)^{d/2} \varrho \exp\left(-\frac{H(W * \varrho)}{G(W * \varrho)^2} |v|^2 \right) \,. \tag{3.10}$$

Therefore, at equilibrium the mean velocity u vanishes if $\varrho \neq 0$, and the "statistical temperature" T, defined by

$$\frac{d}{2} \varrho T = \frac{1}{2} \int_{\mathbb{R}^d} \mathcal{M}[\varrho](v) |v|^2 \, \mathrm{d}v \,,$$

depends non-locally on ϱ and is given by

$$T = T[W * \varrho] = \frac{G(W * \varrho)^2}{2H(W * \varrho)} \,. \tag{3.11}$$

Consequently, in crowded regions (where $W * \varrho$ is high) the temperature is low ("freezing of the aggregates"), while the unhabited regions are "hot".

We introduce the diffusive scaling with the small parameter $\varepsilon > 0$ to (3.5),

$$\varepsilon^2 \frac{\partial f}{\partial t} + \varepsilon v \cdot \nabla_x f = \nabla_v \cdot \left(H(W * \varrho)vf + \frac{1}{2}\nabla_v \left(G(W * \varrho)^2 f \right) \right),$$

and consider the Hilbert expansion in terms of ε, $f = \mathcal{M}[\varrho] + \varepsilon g$, where $\mathcal{M}[\varrho]$ is given by (3.10). Moreover, we perform the Taylor expansion of $H(W * \varrho)$, which we formally write as

$$H(W * \varrho) = H_0 + \varepsilon H_1 + \mathcal{O}(\varepsilon^2),$$

with

$$H_0 = H(W * \varrho_\mathcal{M}), \qquad H_1 = H'(W * \varrho_\mathcal{M})W * \varrho_g,$$

and similarly for $G(W * \varrho) = G_0 + \varepsilon G_1 + \mathcal{O}(\varepsilon^2)$. Here, $\varrho_\mathcal{M}$ and ϱ_g denote the velocity averages of the lowest order terms in the expansion, i.e.

$$\varrho_\mathcal{M} = \int_{\mathbb{R}^d} \mathcal{M}[\varrho]\,\mathrm{d}v, \quad \varrho_g = \int_{\mathbb{R}^d} g\,\mathrm{d}v.$$

Then, collecting terms of order ε, we obtain

$$v \cdot \nabla_x \mathcal{M} = \nabla_v \cdot \left(H_0 vg + H_1 v\mathcal{M} + \frac{1}{2}\nabla_v \left(G_0 g + G_1 \mathcal{M} \right) \right),$$

which yields, after multiplication by v and integration,

$$\frac{d}{2}\nabla_x \left(\varrho_\mathcal{M} T[W * \varrho_\mathcal{M}] \right) = \nabla_x \cdot \int_{\mathbb{R}^d} \mathcal{M}v \otimes v\,\mathrm{d}v = -H_0 \int_{\mathbb{R}^d} gv\,\mathrm{d}v, \tag{3.12}$$

with $T[W * \varrho_\mathcal{M}]$ given by (3.11). Collecting terms of order ε^2 and integrating with respect to v, we obtain

$$\frac{\partial \varrho_\mathcal{M}}{\partial t} + \nabla_x \cdot \int_{\mathbb{R}^d} gv\,\mathrm{d}v = 0,$$

and using (3.12), we finally obtain the nonlinear diffusion equation

$$\frac{\partial \varrho_\mathcal{M}}{\partial t} - \frac{d}{2}\nabla_x \cdot \left(\frac{1}{H(W * \varrho_\mathcal{M})}\nabla_x \left(T[W * \varrho_\mathcal{M}]\varrho_\mathcal{M} \right) \right) = 0. \tag{3.13}$$

Observe that with the choice $H \equiv \text{const.}$, (3.13) reduces to (3.4), possibly up to a linear rescaling of time.

4 Mathematical analysis of the first-order model

In this section we show the existence of weak solutions of the first-order nonlinear diffusion equation (3.4) and some asymptotic properties of these solutions. For simplicity, we consider the full space setting $\Omega = \mathbb{R}^d$, $d \geq 1$, but our analysis can be easily adapted to the case of a bounded domain Ω with homogeneous Neumann boundary conditions, or to the case of periodic boundary conditions, as in the numerical examples of Section 5.

To simplify the notation, we set $F(z) := \frac{1}{2}G(z)^2$, thus we rewrite (3.4) as

$$\frac{\partial \varrho}{\partial t} = \Delta \left(F(W * \varrho)\varrho \right) , \tag{4.1}$$

subject to the initial condition

$$\varrho(t = 0) = \varrho_0 . \tag{4.2}$$

For the rest of this Section, and without further notice, we make the following reasonable assumptions on F and W:

- $F : \mathbb{R}^+ \to \mathbb{R}^+$ is a bounded, nonnegative and nonincreasing function. Let us note that we allow for degeneracy in (3.4), i.e., there might exist an $s_0 > 0$ such that $F(s) = 0$ for all $s > s_0$.

- F is continuously differentiable with globally Lipschitz continuous derivative and $G = \sqrt{2F}$ is globally Lipschitz continuous.

- $W \in W^{1,\infty}(\mathbb{R}^d) \cap H^2(\mathbb{R}^d)$.

Definition 1 *We call $\varrho \in L^\infty(0,T;\mathcal{P}(\Omega))$, where $\mathcal{P}(\Omega)$ denotes the set of probability measures on Ω, a weak solution of (4.1) subject to the initial condition (4.2) with $0 \le \varrho_0 \in L^2(\Omega) \cap \mathcal{P}(\Omega)$, if $\nabla(F(W * \varrho)\varrho) \in L^2(0,T;L^2(\Omega))$ and for every smooth, compactly supported test function $\varphi \in C_c^\infty([0,T) \times \Omega)$ we have*

$$\int_0^\infty \int_\Omega \varrho \frac{\partial \varphi}{\partial t} \, \mathrm{d}x \, \mathrm{d}t + \int_\Omega \varrho_0 \varphi(t = 0) \, \mathrm{d}x = \int_0^\infty \int_\Omega \nabla(F(W * \varrho)\varrho) \cdot \nabla \varphi \, \mathrm{d}x \, \mathrm{d}t . \tag{4.3}$$

The proof of existence of weak solutions in the sense of Definition 1 will be performed in two steps. First, we consider an approximating, uniformly parabolic equation, and prove the existence of its solutions. Then, we remove the approximation in a limiting procedure, to obtain global in time distributional solutions.

4.1 Approximation

In order to obtain a uniformly positive diffusion coefficient, we use the approximation $F_\varepsilon(z) := F(z) + \varepsilon$ for $\varepsilon > 0$, and analyze the approximating equation $\frac{\partial \varrho}{\partial t} = \Delta \left(F_\varepsilon(W * \varrho)\varrho \right)$, which we write in the Fokker-Planck form

$$\frac{\partial \varrho}{\partial t} = \nabla \cdot \left(\varrho \nabla F_\varepsilon(W * \varrho) + F_\varepsilon(W * \varrho)\nabla \varrho \right) \tag{4.4}$$

subject to the inital condition

$$\varrho(t = 0) = \varrho_0 . \tag{4.5}$$

Theorem 1 *For every $\varepsilon > 0$, $T > 0$ and $\varrho_0 \in L^2(\Omega) \cap \mathcal{P}(\Omega)$ there exists a nonnegative weak solution*

$$\varrho^\varepsilon \in L^2(0,T;H^1(\Omega)) \cap H^1(0,T;H^{-1}(\Omega)) \cap L^\infty(0,T;\mathcal{P}(\Omega))$$

of (4.4) in the sense of Definition 1 (with F_ε in place of F).

Proof: The solution is constructed as a fixed point of the map $\Theta : u \mapsto \varrho$, defined on the convex, bounded set

$$\mathcal{B}_R := \{u \in L^2((0,T) \times \Omega)\,;\ \|u\|_{L^\infty(0,T;L^2(\Omega))} \leq R \text{ and } u(t,\cdot) \in \mathcal{P}(\Omega) \text{ for a.e. } t \in (0,T)\}$$

with $R > 0$ to be specified later, where $\varrho = \Theta(u)$ is the unique solution of the Fokker-Planck equation

$$\frac{\partial \varrho}{\partial t} = \nabla \cdot \left(\varrho \nabla F_\varepsilon(W * u) + F_\varepsilon(W * u)\nabla \varrho\right). \tag{4.6}$$

Due to the assumed properties of W and F, the diffusion coefficient satisfies $F_\varepsilon(W * u) \in L^\infty((0,T) \times \Omega)$ and is bounded below by ε, and the convection coefficient $\nabla F_\varepsilon(W * u) \in L^\infty((0,T) \times \Omega)$. Thus, (4.6) is a linear, uniformly parabolic equation with bounded coefficients, and the standard theory [5] implies the existence and uniqueness of the global, nonnegative weak solution

$$\varrho \in L^\infty(0,T;L^2(\Omega)) \cap L^2(0,T;H^1(\Omega)) \cap H^1(0,T;H^{-1}(\Omega)), \tag{4.7}$$

such that $\varrho(\cdot,t) \in \mathcal{P}(\Omega)$ for almost every $t \geq 0$. Therefore, Θ is well defined on $\mathcal{B}_R$ and, as can be easily proven, is continuous there.

In order to apply the Schauder fixed point theorem (cf. [4]), we need to show that Θ maps $\mathcal{B}_R$ into a compact subset of itself. The key point is the following a-priori estimate, obtained by using ϱ as a test function for (4.6) (note that due to (4.7), ϱ is indeed an admissible test function):

$$\int_0^s \int_\Omega \varrho \frac{\partial \varrho}{\partial t}\, \mathrm{d}x\, \mathrm{d}t + \int_\Omega \varrho_0^2\, \mathrm{d}x + \int_0^s \int_\Omega F_\varepsilon |\nabla \varrho|^2\, \mathrm{d}x\, \mathrm{d}t = -\int_0^s \int_\Omega \varrho \nabla F_\varepsilon \cdot \nabla \varrho\, \mathrm{d}x\, \mathrm{d}t,$$

for $s \in (0,T)$, where we introduced the shorthand notation $F_\varepsilon := F_\varepsilon(W * u)$. Now we use the identity

$$\int_0^s \int_\Omega \varrho \frac{\partial \varrho}{\partial t}\, \mathrm{d}x\, \mathrm{d}t = \frac{1}{2} \int_0^s \frac{d}{dt}\left(\int_\Omega \varrho^2\, \mathrm{d}x\right) \mathrm{d}t = \frac{1}{2} \int_\Omega \varrho(x,s)^2\, \mathrm{d}x - \frac{1}{2} \int_\Omega \varrho_0^2\, \mathrm{d}x$$

and the Young inequality

$$\left|\int_0^s \int_\Omega \varrho \nabla F_\varepsilon \cdot \nabla \varrho\, \mathrm{d}x\, \mathrm{d}t\right| \leq \frac{1}{2} \int_0^s \int_\Omega F_\varepsilon |\nabla \varrho|^2\, \mathrm{d}x\, \mathrm{d}t + 2 \int_0^s \int_\Omega \left|\nabla \sqrt{F_\varepsilon}\right|^2 \varrho^2\, \mathrm{d}x\, \mathrm{d}t$$

to obtain

$$\int_\Omega \varrho(x,s)^2\, \mathrm{d}x + \frac{1}{2} \int_0^s \int_\Omega F_\varepsilon |\nabla \varrho|^2\, \mathrm{d}x\, \mathrm{d}t \leq \int_\Omega \varrho_0^2\, \mathrm{d}x + \int_0^s \int_\Omega \left|\nabla \sqrt{F_\varepsilon}\right|^2 \varrho^2\, \mathrm{d}x\, \mathrm{d}t.$$

Finally, since $F_\varepsilon \geq \varepsilon$ and $|F'(y)| \leq C$ for $0 \leq y \leq \sup_{x \in \Omega} W * u(x) \leq \|W\|_{L^\infty}$,

$$\left|\nabla \sqrt{F_\varepsilon}\right|^2 = \left|\nabla \sqrt{F_\varepsilon(W * u)}\right|^2 = \frac{|F'(W * u)|^2}{4F_\varepsilon(W * u)}|\nabla W * u|^2 \leq \frac{C^2}{4\varepsilon} \|W\|_{W^{1,\infty}}^2 =: C_\varepsilon,$$

and we conclude

$$\int_\Omega \varrho(x,s)^2\, \mathrm{d}x + \frac{\varepsilon}{2} \int_0^s \int_\Omega |\nabla \varrho|^2\, \mathrm{d}x\, \mathrm{d}t \leq \int_\Omega \varrho_0^2\, \mathrm{d}x + C_\varepsilon \int_0^s \int_\Omega \varrho^2\, \mathrm{d}x\, \mathrm{d}t.$$

The Gronwall inequality yields an a-priori bound on ϱ in $L^\infty(0,T;L^2(\Omega))$, which implies that Θ indeed maps $\mathcal{B}_R$ into itself, for a suitable $R > 0$. Moreover, reinserting into the right-hand side, we obtain an a-priori bound in $L^2(0,T;H^1(\Omega))$, and an analysis of the right-hand side of (4.6) immediately yields a bound

in $H^1(0, T; H^{-1}(\Omega))$. The proof is concluded by an application of the Aubin-Lions Lemma (cf. [4]), which states that the space $L^2(0, T; H^1(\Omega)) \cap H^1(0, T; H^{-1}(\Omega))$ is compactly embedded into $L^2((0, T) \times \Omega)$, so that Θ maps $\mathcal{B}_R$ into a compact subset of itself.

∎

Most a-priori estimates on ϱ^ε that have been obtained along the lines of the above proof are not uniform in ε and can thus not be used to pass to the limit $\varepsilon \to 0$. However, we can tackle this problem with a more careful analysis, where the key idea is to look for estimates on $\sqrt{F_\varepsilon} \varrho^\varepsilon$, rather than on ϱ^ε.

Lemma 1 *There exists a constant C independent of $\varepsilon > 0$ small enough, such that the solutions ϱ^ε of (4.4), constructed in Theorem 1, satisfy*

$$\|\varrho^\varepsilon\|_{L^\infty(0,T;\mathcal{P}(\Omega))} \leq 1, \qquad \|\varrho^\varepsilon\|_{L^\infty(0,T;L^2(\Omega))} \leq C, \qquad \left\| \frac{\partial \varrho^\varepsilon}{\partial t} \right\|_{L^2(0,T;H^{-1}(\Omega))} \leq C,$$

and

$$\|W * \varrho^\varepsilon\|_{L^\infty(0,T;W^{1,\infty}(\Omega))} \leq C, \qquad \left\| W * \frac{\partial \varrho^\varepsilon}{\partial t} \right\|_{L^\infty(0,T;L^2(\Omega))} \leq C.$$

Proof: For the sake of better legibility, we will drop the superscript at ϱ^ε in the proof. Since ϱ is constructed as a nonnegative weak solution of the Fokker-Planck equation (4.4), the first estimate follows immediately due to the mass conservation

$$\|\varrho(t, \cdot)\|_{L^1(\Omega)} = \int_\Omega \varrho(t, x) \, dx = \int_\Omega \varrho_0(x) \, dx = 1.$$

Proceeding in a similar way as in the proof of Theorem 1, we conclude for $s \in (0, T)$

$$\frac{1}{2} \int_\Omega \varrho(s, x)^2 \, dx + \int_0^s \int_\Omega F_\varepsilon |\nabla \varrho|^2 \, dx \, dt + \int_0^s \int_\Omega \varrho \nabla F_\varepsilon \cdot \nabla \varrho \, dx \, dt = \int_\Omega \varrho_0^2 \, dx,$$

where we again use the shorthand notation $F_\varepsilon := F_\varepsilon(W * \varrho)$. Using the identity

$$\varrho \nabla F_\varepsilon \cdot \nabla \varrho + F_\varepsilon |\nabla \varrho|^2 = \left| \nabla \left(\sqrt{F_\varepsilon} \varrho \right) \right|^2 - \left| \varrho \nabla \sqrt{F_\varepsilon} \right|^2,$$

we obtain

$$\frac{1}{2} \int_\Omega \varrho(s, x)^2 \, dx + \int_0^s \int_\Omega \left| \nabla \left(\sqrt{F_\varepsilon} \varrho \right) \right|^2 \, dx \, dt = \frac{1}{2} \int_\Omega \varrho_0^2 \, dx + \int_0^s \int_\Omega \varrho^2 \left| \nabla \sqrt{F_\varepsilon} \right|^2 \, dx \, dt.$$

Now, since $|G'(y)| \leq C$ for $0 \leq y \leq \sup_{x \in \Omega} W * u(x) \leq \|W\|_{L^\infty}$, we have

$$\left| \nabla \sqrt{F_\varepsilon} \right| = \left| \nabla \sqrt{F_\varepsilon(W * \varrho)} \right| = |G'(W * \varrho)||\nabla W * \varrho| \leq C \|W\|_{W^{1,\infty}}.$$

Hence,

$$\frac{1}{2} \int_\Omega \varrho(s, x)^2 \, dx + \int_0^s \int_\Omega \left| \nabla \left(\sqrt{F_\varepsilon} \varrho \right) \right|^2 \, dx \, dt \leq \frac{1}{2} \int_\Omega \varrho_0^2 \, dx + C^2 \|W\|_{W^{1,\infty}}^2 \int_0^s \int_\Omega \varrho^2 \, dx \, dt,$$

12

and the Gronwall inequality yields a uniform estimate for ϱ in $L^\infty(0,T;L^2(\Omega))$, which subsequently implies a uniform estimate for $\sqrt{F_\varepsilon(W*\varrho)}\varrho$ in $L^2(0,T;H^1(\Omega))$ and, subsequently, also for $\frac{\partial\varrho}{\partial t}$ in $L^2(0,T;H^{-1}(\Omega))$.

Finally, we derive the estimates for $W*\varrho$. Since $W \in W^{1,\infty}(\mathbb{R}^d)$, we immediately obtain

$$\|W*\varrho\|_{L^\infty(0,T;W^{1,\infty}(\Omega))} \le \|W\|_{W^{1,\infty}(\mathbb{R}^d)} \sup_{t\in[0,T]} \int_\Omega |\varrho(t,x)|\,\mathrm{d}x = \|W\|_{W^{1,\infty}(\mathbb{R}^d)}\,.$$

Moreover, we have

$$W*\frac{\partial\varrho}{\partial t} = W*\Delta(F_\varepsilon\varrho) = \Delta W*(F_\varepsilon\varrho)\,,$$

and with $W \in H^2(\mathbb{R}^d)$, $\varrho \in L^\infty(0,T;\mathcal{P}(\Omega))$ and the uniform boundedness of F_ε in $L^\infty((0,T)\times\Omega)$, we obtain a uniform bound for $W*(\frac{\partial\varrho}{\partial t})$ in $L^\infty(0,T;L^2(\Omega))$. ∎

4.2 Global existence

From the above approximation and a-priori estimates we can easily pass to the limit $\varepsilon \to 0$ and conclude the existence of a weak solution for (4.1):

Theorem 2 *Let $\varrho_0 \in L^2(\Omega)\cap\mathcal{P}(\Omega)$ and $T > 0$. Then there exists a solution*

$$\varrho \in L^\infty(0,T;L^2(\Omega)) \cap L^\infty(0,T;\mathcal{P}(\Omega)) \cap H^1(0,T;H^{-1}(\Omega))$$

of (4.1)–(4.2) in the sense of the weak formulation (4.3), such that in addition

$$\sqrt{F(W*\varrho)}\varrho \in L^2(0,T;H^1(\Omega))\,.$$

Proof: We use the uniform estimates of Lemma 1 and the Banach-Alaoglu theorem [4] to extract a subsequence $\varrho^{\varepsilon n}$ such that

$$
\begin{aligned}
\varrho^{\varepsilon n} &\rightharpoonup^* \varrho & &\text{weakly-* in } L^\infty(0,T;L^2(\Omega))\,,\\
\frac{\partial\varrho^{\varepsilon n}}{\partial t} &\rightharpoonup \frac{\partial\varrho}{\partial t} & &\text{weakly in } L^2(0,T;H^{-1}(\Omega))\,,\\
\sqrt{F_{\varepsilon n}(W*\varrho^{\varepsilon n})}\varrho^{\varepsilon n} &\rightharpoonup u & &\text{weakly in } L^2(0,T;H^1(\Omega))\,,
\end{aligned}
$$

for some $u \in L^2(0,T;H^1(\Omega))$. Moroever, we use a variant of the Aubin-Lions Lemma [28] to conclude the compact embedding of $L^\infty(0,T;W^{1,\infty}(\Omega))\cap W^{1,\infty}(0,T;L^2(\Omega))$ into $L^\infty(0,T;C(\Omega))$. Thus, we can extract a further subsequence, again denoted by $\varrho^{\varepsilon n}$, such that as $\varepsilon_n \to 0$,

$$W*\varrho^{\varepsilon n} \to v \quad \text{strongly in } L^\infty(0,T;C(\Omega))\,.$$

Since further

$$W*\varrho^{\varepsilon n} \rightharpoonup^* W*\varrho \quad \text{weakly-* in } L^\infty(0,T;L^2(\Omega))\,,$$

13

we conclude $v = W * \varrho$ by the uniqueness of the limit. Due to the continuity properties of F and F', we also have

$$\begin{aligned}
F_{\varepsilon_n}(W * \varrho^{\varepsilon_n}) &\to F(W * \varrho) && \text{strongly in } L^\infty(0, T; C(\Omega)), \\
\sqrt{F_{\varepsilon_n}(W * \varrho^{\varepsilon_n})} &\to \sqrt{F(W * \varrho)} && \text{strongly in } L^\infty(0, T; C(\Omega)), \\
F'_{\varepsilon_n}(W * \varrho^{\varepsilon_n}) &\to F'(W * \varrho) && \text{strongly in } L^\infty(0, T; C(\Omega)).
\end{aligned}$$

Consequently,

$$\sqrt{F_{\varepsilon_n}(W * \varrho^{\varepsilon_n})}\varrho^{\varepsilon_n} \rightharpoonup \sqrt{F(W * \varrho)}\varrho \qquad \text{weakly-* in } L^\infty(0, T; L^2(\Omega)),$$

and, again, by the uniqueness of the limit we identify $u = \sqrt{F(W * \varrho)}\varrho$. Finally, we use the identity

$$\nabla\left(F_{\varepsilon_n}\left(W * \varrho^{\varepsilon_n}\right)\varrho^{\varepsilon_n}\right) = F'_{\varepsilon_n}\left(W * \varrho^{\varepsilon_n}\right)\varrho^{\varepsilon_n}\nabla\left(W * \varrho^{\varepsilon_n}\right) + \sqrt{F_{\varepsilon_n}\left(W * \varrho^{\varepsilon_n}\right)}\nabla\left(\sqrt{F_{\varepsilon_n}\left(W * \varrho^{\varepsilon_n}\right)}\varrho^{\varepsilon_n}\right)$$

and the above limits to conclude

$$\nabla\left(F_{\varepsilon_n}\left(W * \varrho^{\varepsilon_n}\right)\varrho^{\varepsilon_n}\right) \rightharpoonup \nabla\left(F\left(W * \varrho\right)\varrho\right) \qquad \text{weakly in } L^2(0, T; L^2(\Omega)).$$

Hence, we can pass to the limit $\varepsilon_n \to 0$ in the weak formulation (4.3) and conclude that ϱ is a weak solution of (4.1)–(4.2).

■

Finally, we want to reduce the regularity of the initial value towards solely probability measures. This is relevant since we shall see below that indeed Dirac δ distributions can be stationary solutions. For this sake we define the very weak (distributional) notion of the solution:

Definition 2 *We call $\varrho \in L^\infty(0, T; \mathcal{P}(\Omega))$ a very weak (distributional) solution of (4.1) subject to the initial condition (4.2) with $\varrho_0 \in \mathcal{P}(\Omega)$, if for every smooth, compactly supported test function $\varphi \in C_c^\infty([0, T) \times \Omega)$ we have*

$$\int_0^\infty \int_\Omega \frac{\partial\varphi}{\partial t}\varrho \, dx \, dt + \int_\Omega \varphi(t = 0)\varrho_0 \, dx = -\int_0^\infty \int_\Omega (F(W * \varrho)\Delta\varphi)\varrho \, dx \, dt, \tag{4.8}$$

where we denote by $\varrho \, dx$ the integration with respect to the probability measure $\varrho(t, \cdot)$.

Theorem 3 *Let $\varrho_0 \in \mathcal{P}(\Omega)$ and $T > 0$. Then there exists a very weak solution*

$$\varrho \in L^\infty(0, T; \mathcal{P}(\Omega))$$

of (4.1)–(4.2) in the sense of (4.8).

Proof: Let us consider a sequence $(\varrho_0^n)_{n\in\mathbb{N}} \subseteq L^2(\Omega) \cap \mathcal{P}(\Omega)$ such that $\varrho_0^n \to \varrho_0$ in $\mathcal{P}(\Omega)$. Moreover, let us denote by ϱ^n the corresponding weak solutions of (4.1)–(4.2), which are trivially also the very weak solutions in the sense of (4.8). Conservation of mass provides immediately a uniform bound on ϱ^n in $L^\infty(0, T; \mathcal{P}(\Omega))$, such that for a subsequence, again denoted by ϱ^n, we have

$$\varrho^n \rightharpoonup^* \varrho \quad \text{weakly-* in } L^\infty(0, T; \mathcal{P}(\Omega)).$$

14

To show that ϱ is a solution to (4.8), we need to prove that $F(W * \varrho^n)$ converges to $F(W * \varrho)$ strongly in $L^1(0, T; C(\Omega))$. First, let us note that, due to the assumptions on W, $W * \varrho^n$ is uniformly bounded in $L^\infty(0, T; W^{1,\infty}(\Omega))$, and the same holds for $F(W * \varrho^n)$. Consequently, the sequence $F(W * \varrho^n)\varrho^n$ is uniformly bounded in $L^\infty(0, T; \mathcal{P}(\Omega))$. From (4.1) it immediately follows that $\frac{\partial \varrho^n}{\partial t} = \Delta(F(W * \varrho^n)\varrho^n)$ is uniformly bounded in $L^\infty(0, T; (C_c^2(\Omega))^*)$, where $(C_c^2(\Omega))^*$ denotes the dual space to $C_c^2(\Omega)$, the space of twice continuously differentiable functions with compact support in Ω. Consequently, the sequence $\frac{\partial}{\partial t} F(W * \varrho^n) = F'(W * \varrho^n) W * \frac{\partial \varrho^n}{\partial t}$ is uniformly bounded in $L^\infty(0, T; (C_c^2(\Omega))^*)$, and the generalization of the Aubin-Lions lemma by Simon [28] implies then the strong convergence of $F(W * \varrho^n)$ to $F(W * \varrho)$ in $L^1(0, T; C(\Omega))$.

∎

4.3 Stationary Solutions

The numerical simulations provided in Section 5 suggest that for $G > 0$ (for instance, $G(s) = e^{-s}$) and bounded domain with periodic boundary conditions, the stationary solutions of (3.4) consist of one or more localized, but not compactly supported aggregates (clumps), see the bottom right panel of Fig. 5 and Fig. 7. However, we were not able to characterize these stationary aggregates analytically. Instead, we provide a few examples of other types of stationary solutions, posed either in the full space setting $\Omega = \mathbb{R}^d$ or on a torus $\Omega = \mathbb{T}^d$ with periodically extended W. These examples are rather trivial, however, still provide an interesting insight into the relatively rich structure of the solutions of (3.4).

- The most trivial type of stationary solution is the constant state $\varrho \equiv c > 0$. Clearly, this has finite mass only if $\Omega = \mathbb{T}^d$.

- If there exists an $s_0 > 0$ such that $G(s) \equiv 0$ for all $s \geq s_0$, then any profile ϱ such that $\varrho \geq s_0 \left(\int_\Omega W(x)\,dx \right)^{-1}$ almost everywhere on Ω is a stationary solution to the distributional formulation (4.8). Indeed, such a solution satisfies $W * \varrho \geq s_0$, so that $G(W * \varrho) \equiv 0$. However, again, this solution has infinite mass if $\Omega = \mathbb{R}^d$.

- If there exists an $s_0 > 0$ such that $G(s) \equiv 0$ for all $s \geq s_0$, and, moreover, W is continuous on Ω, we construct the atomic measure

$$\varrho(x) = \sum_{i=1}^N c_i \delta(x - x_i)$$

for some $N \in \mathbb{N}$, $x_i \in \Omega$ and $c_i > 0$ such that $c_i W(0) > s_0$ for all $i = 1, \ldots, N$. Then ϱ is a distributional stationary solution to (4.8). Indeed, for any $i = 1, \ldots, N$ we have

$$W * \varrho(x_i) = \sum_{j=1}^N c_j W(x_i - x_j) \geq c_i W(0) > s_0 \,.$$

By the continuity of W and, consequently, $W * \varrho$, we have $W * \varrho > s_0$ on some neighbourhood of x_i. Therefore, $G(W * \varrho) \equiv 0$ on some open set containing $\bigcup_{i=1}^N x_i$ and, finally, $G(W * \varrho)\varrho \equiv 0$ everywhere on Ω.

4.4 Linear stability analysis

In this section we perform a linear stability analysis of constant density states of the nonlinear diffusion equation (3.4). As discussed previously, we have constant steady state solutions either in the full-space setting $\Omega = \mathbb{R}^d$ (however, with infinite mass), or on a bounded domain $\Omega \subset \mathbb{R}^d$ with periodic boundary conditions. Without loss of generality, we assume the normalization $\int_\Omega W(x)\,dx = 1$. Let us make the perturbation ansatz $\varrho = \varrho_0 + \varepsilon\tilde{\varrho}$, where $\varrho_0 > 0$ is a constant state such that $G(\varrho_0) > 0$ and $G'(\varrho_0) < 0$. Then $W * \varrho = \varrho_0 + \varepsilon W * \tilde{\varrho}$ and assuming sufficient smoothness of G, we have

$$G(W * \varrho)^2 = G(\varrho_0)^2 + 2\varepsilon G(\varrho_0)G'(\varrho_0)W * \tilde{\varrho} + O(\varepsilon^2)\,.$$

Inserting the ansatz into (3.4) and collecting terms of order $\mathcal{O}(\varepsilon)$, we arrive at

$$\frac{\partial\tilde{\varrho}}{\partial t} - \frac{G(\varrho_0)}{2}\left(G(\varrho_0)\Delta\tilde{\varrho} + 2G'(\varrho_0)\varrho_0\Delta(W * \tilde{\varrho})\right) = 0\,.$$

Performing the Fourier transform, we obtain

$$\frac{\partial\hat{\varrho}}{\partial t} + |\xi|^2\frac{G(\varrho_0)}{2}\left(G(\varrho_0) + 2G'(\varrho_0)\varrho_0\hat{W}\right)\hat{\varrho} = 0\,, \tag{4.9}$$

where we denoted $\hat{\varrho} = \hat{\varrho}(t,\xi)$ the Fourier transform of $\tilde{\varrho}$. Consequently, with the assumption $G(\varrho_0) > 0$ and $G'(\varrho_0) < 0$, those wavenumbers ξ of $\tilde{\varrho}$ are stable for which

$$\mathrm{Re}\,\hat{W}(\xi) < -\frac{G(\varrho_0)}{2G'(\varrho_0)\varrho_0} \geq 0\,. \tag{4.10}$$

Since $W \in L^1(\mathbb{R}^d)$, we have $\hat{W} \in C^0(\mathbb{R}^d)$, and, therefore, all wavenumbers of $\tilde{\varrho}$ larger than a certain threshold will be stable. On the other hand, wavenumbers violating (4.10) will lead to pattern formation, as we show in our numerical examples, Section 5. Moreover, a quick inspection of (4.9) leads to the expectation that the stable high wavenumbers will be smoothed-out on a faster time scale, compared to the slower emergence of patterns due to the unstable lower wavenumbers. Finally, considering the scalings $W_r(x) := r^{-d}W_1(x/r)$ of a fixed kernel $W_1 \in L^1(\mathbb{R}^d)$, we have $\hat{W}_r(\xi) = \hat{W}_1(r\xi)$ and, therefore, we expect that the wavelength of the patterns (size of aggregates) will scale with r. This can be as well clearly observed in our numerical examples.

It is interesting to consider the formal extremal case $W = \delta_0$, i.e., $W * \tilde{\varrho} = \tilde{\varrho}$. Then $\hat{W} \equiv 1$ and we conclude that the constant state ϱ_0 is stable if and only if

$$(G(\varrho_0)^2\varrho_0)' = G(\varrho_0)^2 + 2G(\varrho_0)G'(\varrho_0)\varrho_0 > 0\,. \tag{4.11}$$

In fact, violation of (4.11) with $W = \delta_0$ leads to an ill-posed problem, since then (3.4) looks like a backward heat equation, which can be seen if we expand the derivatives and write it in the Fokker-Planck form as

$$\frac{\partial\varrho}{\partial t} = \frac{1}{2}\nabla\cdot\left(G(\varrho)[2G'(\varrho)\varrho + G(\varrho)]\nabla\varrho\right)\,. \tag{4.12}$$

The equation is parabolic (and thus well-posed) only if the diffusivity $2G(\varrho)G'(\varrho)\varrho + G^2(\varrho)$ is strictly positive, which is our stability condition (4.11). Therefore, only imposing an initial condition ϱ_0 uniformly satisfying (4.11) leads to a well-posed diffusion equation for all $t \geq 0$, since (4.11) is preserved due to

16

the maximum principle. On the other hand, if $G > 0$, the nonlocality $W \in L^1(\mathbb{R}^d)$ always stabilizes the equation in the sense of (4.10). Indeed, writing it in the Fokker-Planck form

$$\frac{\partial \varrho}{\partial t} = \frac{1}{2} \nabla \cdot \left(Individual based and mean-field modelling of direct aggregation \left[2G(W*\varrho)G'(W*\varrho)\nabla W*\varrho \right] \varrho + G^2(W$$

we see that the second-order term appears with the positive diffusivity $G^2(W*\varrho)$. The first-order term describes then the transport of ϱ along the generalized gradient $\nabla W*\varrho$ and is responsible for the aggregative effect.

Clearly, the ill-posedness of (4.12) can be avoided by merely introducing a nonlocality in the first-order transport term, while the diffusivity may stay local (and possibly degenerate). Such a model was derived and studied in [30], which with our notation is written as

$$\frac{\partial \varrho}{\partial t} = \nabla \cdot \left(-\varrho \nabla W * \varrho + \varrho^2 \nabla \varrho \right) .$$

This equation was constructed as a model for biological aggregations in which individuals experience long-range social attraction and short range dispersal. Let us note that here, in contrast to our model, the diffusivity is an increasing function of the density ϱ. In [30] it was shown that it produces strongly nonlinear states with compact support and steep edges that correspond to localized biological aggregations, or clumps. Similarly as can be observed in our numerical simulations in Section 5, these clumps are approached through a dynamic coarsening process.

Another insight into the stabilizing effect of the nonlocality is provided by the introduction of a formal expansion of W. Taking W as the standard mollifier and $W_\varepsilon(s) := \varepsilon^{-d}W(s/\varepsilon)$, such that $W_\varepsilon \to \delta_0$ as $\varepsilon \to 0$, we expand

$$
\begin{aligned}
W_\varepsilon * \varrho &= \varepsilon^{-d} \int_{\mathbb{R}^d} W\left(\frac{x-y}{\varepsilon}\right) \varrho(y)\,\mathrm{d}y \\
&= \int_{\mathbb{R}^d} W(-z)\varrho(x+\varepsilon z)\,\mathrm{d}z \\
&= \int_{\mathbb{R}^d} W(-z)\left(\varrho(x) + \varepsilon \nabla \varrho(x)z + \frac{1}{2}\varepsilon^2 z^t \nabla^2 \varrho(x)z \right)\,\mathrm{d}z + \mathcal{O}(\varepsilon^3) .
\end{aligned}
$$

Now, due to the symmetry $W(-z) = W(z)$ and the normalization $\int_{\mathbb{R}^d} W(z)\,\mathrm{d}z = 1$, we have

$$W_\varepsilon * \varrho = \varrho + \frac{\varepsilon^2}{2}\beta\Delta\varrho + \mathcal{O}(\varepsilon^4) ,$$

where the constant $\beta > 0$ is such that $\int_{\mathbb{R}^d} W(z)z_i z_j\,\mathrm{d}z = \beta\delta_{ij}$. Inserting this into (3.4), we obtain

$$
\begin{aligned}
\partial_t \varrho &= \frac{1}{2}\Delta\left(G\left(\varrho + \frac{\varepsilon^2}{2}\beta\Delta\varrho + \mathcal{O}(\varepsilon^4) \right)^2 \varrho \right) \\
&= \frac{1}{2}\Delta\left(G(\varrho)^2\varrho + \varepsilon^2\beta G(\varrho)G'(\varrho)\varrho\Delta\varrho \right) + \mathcal{O}(\varepsilon^4) .
\end{aligned}
$$

Up to the terms of fourth order in ε, this is a Cahn-Hilliard-type equation which is well-posed for every $\varepsilon > 0$ since the term $\varepsilon^2\beta G(\varrho)G'(\varrho)\varrho$ is strictly negative if $G(\varrho) > 0$ and $G'(\varrho) < 0$. However, if $\varepsilon = 0$, i.e. $W = \delta_0$, this regularizing effect is lost.
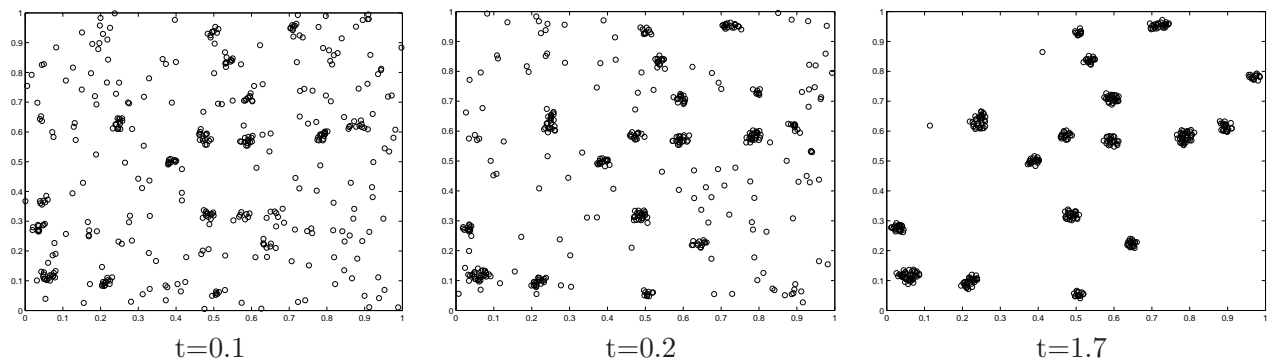
17
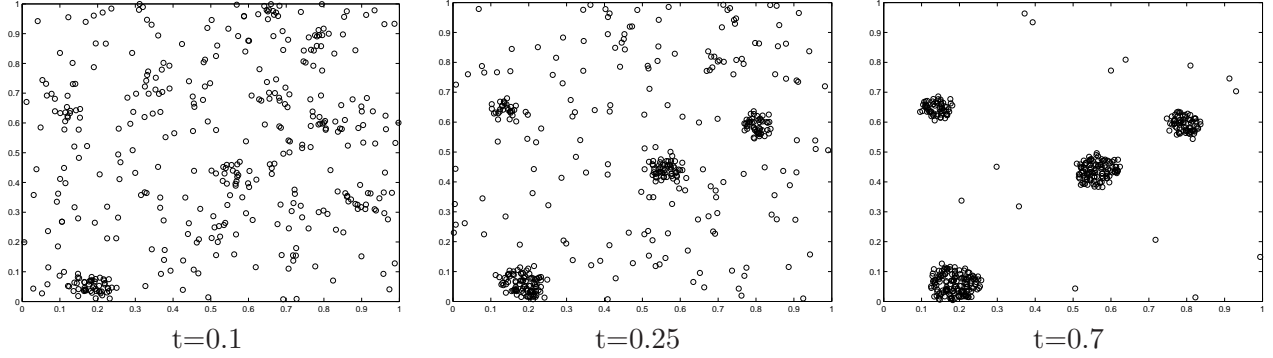
Figure 1: The first-order individual based model with $N = 400$ agents and sampling radius $R = 0.025$, subject to a random initial condition.

# 5 Numerical examples

In this section we present numerical examples for the first and second-order individual based models (2.2), (2.4)–(2.5) in 2D, the first-order mean-field limit (3.4) in 1D and 2D and the second-order mean-field limit (3.5) in 1D with 1D velocity.

5.1 The first-order individual based model (2.2)

We consider a system consisting of $N = 400$ individuals in a 2D domain $\Omega = (0,1) \times (0,1)$ with periodic boundary conditions. The initial positions are generated randomly and independently for each individual from the uniform distribution on Ω; we took the same initial condition for all the three experiments below. We choose $W(x) = w(|x|)$ with w the characteristic function of the interval $[0, R]$, corresponding to the sampling radius R, and for R we take the values 0.025 (Fig. 1), 0.05 (Fig. 2) and 0.1 (Fig. 3). For G we make the choice $G(s) = \exp(-s/3)$. The system of stochastic differential equations (2.2) is integrated in time using the Euler-Mayurama scheme with time-step length $t = 10^{-3}$. We used the linear stability analysis in Section 5 to make the "right" choice of G, such that we could observe pattern formation. Indeed, if G decreases too quickly, the system will "freeze" immediately, before any aggregates could be formed; on the other hand, if G does not decrease fast enough, the system is "overheated" and does not allow aggregates to persist.

In Figs. 1, 2 and 3 we observe that with a smaller sampling radius R, a larger number of small aggregates is created on a faster time scale. The choice $R = 0.1$ (Fig.3) leads to creation of one sigle aggregate. This aggregate is approximately ring-shaped, with higher density of particles around the circumference and lower density in the middle. This can be explained by the fact that the aggregate grows by "capturing" particles from its neighborhood, and once a particle is captured, its mobility is greatly reduced, so that it only slowly makes its way towards the center of the aggregate. Let us also mention that the right-most panels in Figs. 1–3 present quasi-steady states, where the aggregates are in a dynamic equilibrium with a very few free-running particles. However, on a very long time scale, the smaller aggregates typically disintegrate and their particles are caught by the larger ones. Such a coarsening behavior is typical to diffusive aggregation systems, as for instance the Cahn-Hilliard equation, or the nonlocal continuum model of [30].

Figure 2: The first-order individual based model with $N = 400$ agents and sampling radius $R = 0.05$, subject to a random initial condition.
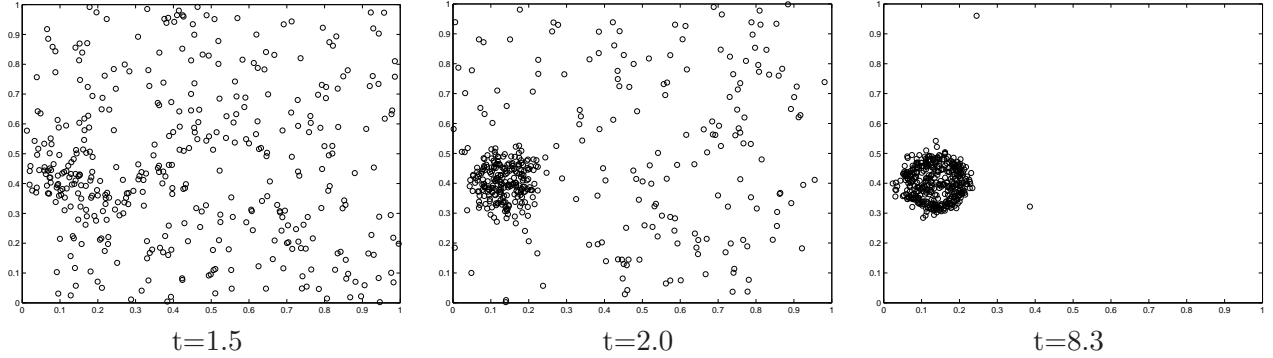
Figure 3: The first-order individual based model with $N = 400$ agents and sampling radius $R = 0.1$, subject to a random initial condition.

5.2 The second-order individual based model (2.4)–(2.5)

Generally, the behavior of the second-order individual based model is very similar to this of the first-order one, at least if we do not impose a restricted cone of vision (i.e., W in (2.3) does not depend on v). Indeed, with a suitable choice of parameters, one again observes the formation of quasi-stable aggregates, whose number and size depend on the sampling radius. The most striking difference with respect to the first-order model is that the movement of the agents is smoother (their velocities are continuous) and the shape of the aggregates is more complex (we observed emergence of ellipsoidal aggregates, instead of the almost-circular ones in the first-order model).

The situation becomes slightly more interesting if we consider a restricted cone of vision - we choose 180°, such that $W(x,v) = w\left(|x|, \frac{x \cdot v}{|x||v|}\right)$ with $w(s,z) = \chi_{[0,R]}(s)\chi_{[0,1]}(z)$. The sampling radius is chosen as $R = 0.05$. Again, we simulate with $N = 400$ individuals in a 2D domain $\Omega = (0,1) \times (0,1)$ with periodic boundary conditions, and $G(s) = \exp(-s/3)$. For H we set the constant $H \equiv 2$. The system of stochastic

19
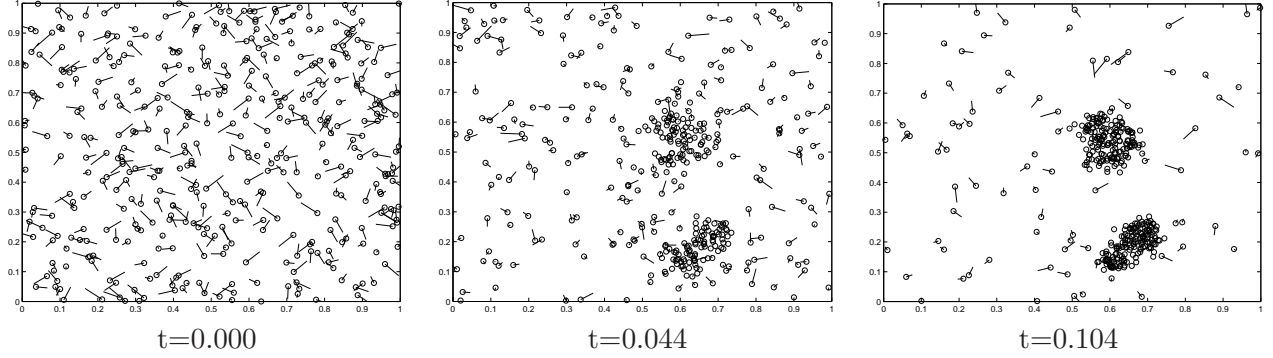
Figure 4: The second-order individual based model with $N = 400$ agents, sampling radius $R = 0.05$ and 180° cone of vision, subject to a random initial condition.

differential equations (2.4)–(2.5) is integrated in time using the Euler-Mayurama scheme with time-step length $t = 10^{-3}$. The initial positions of the agents are generated randomly in the same way as for the first-order model, while their initial velocities are generated independently and randomly from the 2D normalized Gaussian distribution. Three snapshots of the evolution of the system are shown in Fig. 4, the velocities being marked by the linear elements for each individual. The right-most panel in Fig. 4 shows the quasi-steady state with two aggregates formed, in dynamic equilibrium with the "free-running" individuals. It is obvious that, compared to the previous simulations, the aggregates are less densely packed and their shapes are less circular. Also, the portion of the "free-running" particles is much higher, which clearly is an effect of the restricted cone of vision (the captured particles can leave the aggregate more easily).

5.3 The first-order mean-field model (3.4) in 1D

We simulated the first-order mean-field model (3.4) in the 1D periodic domain $(0, 1)$, using semi-implicit finite difference discretization for the space variable and first-order forward Euler method for the time variable. The space grid consisted of 200 equidistant points, the time step was 10^{-4}. As before, we chose $W(x) = w(|x|)$ with w the characteristic function of the interval $[0, 0.1]$, and $G(s) = \exp(-s/3)$. We imposed a random initial condition for ϱ, generated such that for every grid point a random number from the uniform distribution in $[0, 1]$ has been drawn. Snapshots of the evolution are shown in Fig. 5. We observe that quite a strong smoothing effect takes place on the fast time-scale (first row in Fig. 5), while aggregation takes place on a time-scale approximately one order of magnitude slower (second row); this is explained with the stability analysis in Section 4.4. First, two aggregates of different sizes are created, however, both of them are unstable - the smaller one is smoothed out, while the larger grows further, until the steady state is reached (lower right panel). Observe also the characteristic "fork"-like shape of the profile in the lower mid panel ($t = 2.65$), which is due to the mass arriving from the neighbourhood with higher diffusivity than the diffusivity in the middle of the profile; compare also with the ring-shaped aggregate in the right panel of Fig. 2. However, this fork-like structure is eventually also smoothed out, to finally obtain the steady profile. Let us note that the steady aggregate, although well localized, is not compactly supported, i.e., the profile has positive density everywhere on $[0, 1]$.
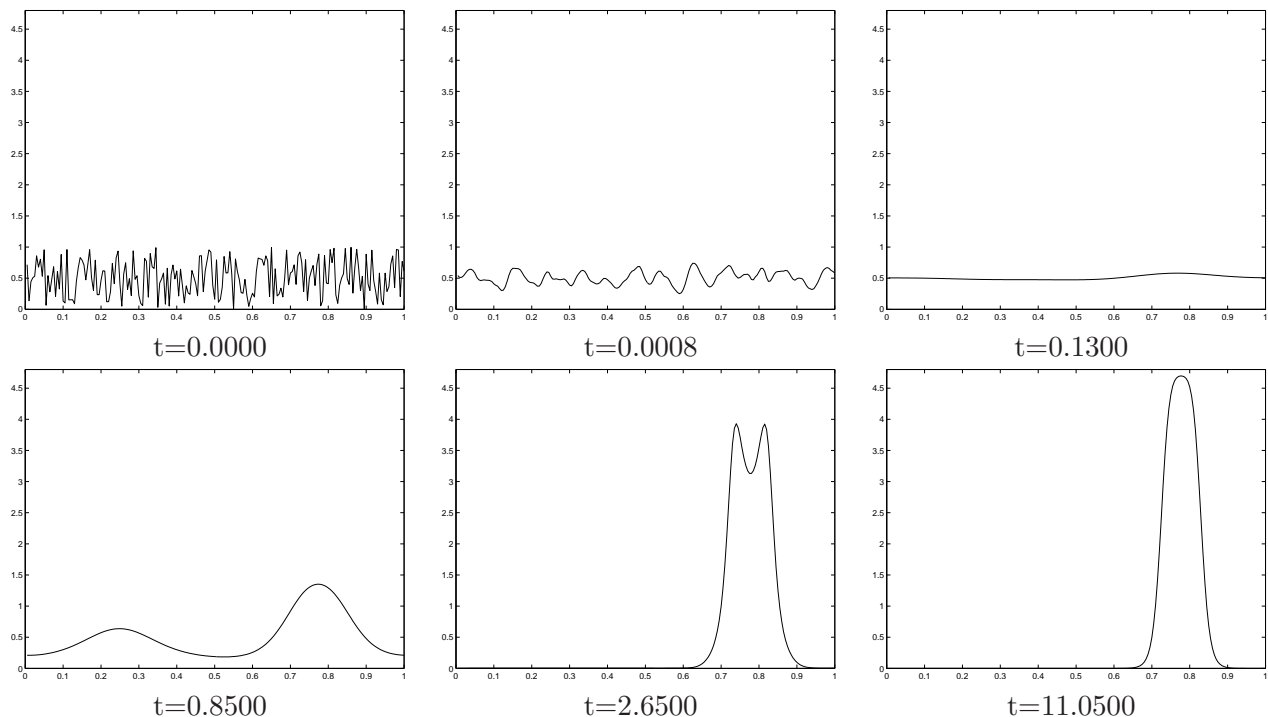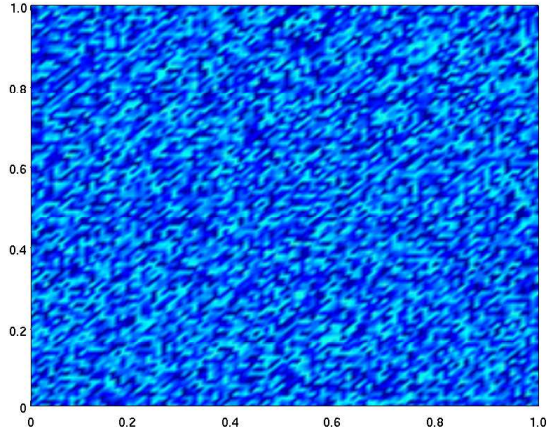
20

Figure 5: The first-order mean-field model in a periodic 1D setting with random initial condition (upper left panel). Steady state was reached at $t \simeq 11$ (lower right panel).
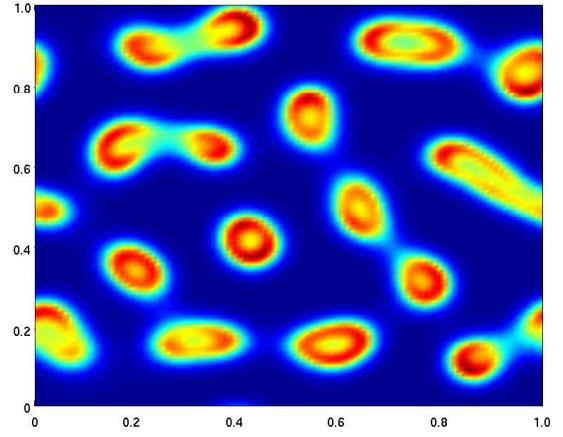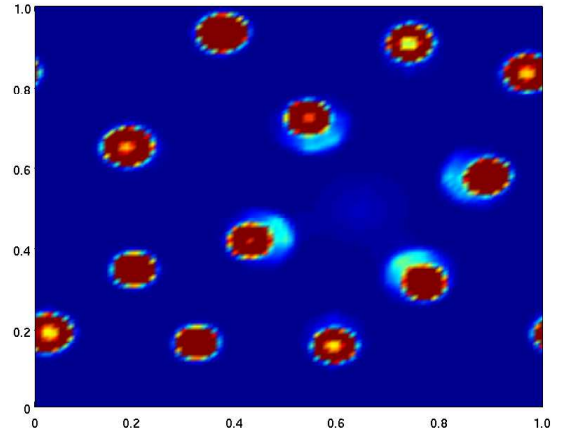
5.4 The first-order mean-field model (3.4) in 2D

We simulated the first-order mean-field model (3.4) in the 2D periodic domain $(0,1)^2$, using the same type of discretization as in the 1D case. The space grid consisted of 100×100 equidistant points, the time step was 10^{-4}. We chose the sampling radius $R = 0.07$, i.e., $W(x) = \chi_{[0,0.7]}(|x|)$, and $G(s) = \exp(-s/3)$. Snapshots of the evolution are shown in Fig. 6. Starting again from a random initial condition, we observed a rapid formation of approximately ring-shaped pre-aggregates, which eventually turn into an almost regular pattern of well localized (but not compactly supported) clumps. However, the smaller aggregates may be unstable and diffusively disintegrate, and their mass is absorbed by their neighbors, as shown on the lower right panel of Fig. 6. Finally, in Fig. 7 we present examples of patterns produced with the sampling radii $R = 0.6$ (left panel) and $R = 0.11$ (right panel).

5.5 The second-order mean-field model (3.5)

We conclude this section with simulations of the second-order mean-field model (3.5) in 1D. The spatial domain $\Omega = [a,b]$ is discretized using an equidistant mesh $x_i = a + i\Delta x$, the velocity domain $V = [v_{\min}, v_{\max}]$ at grid points $v_j = v_{\min} + j\Delta v$. The time step Δt satisfies the CFL condition for $v_{\max}$, i.e., $\Delta t = \frac{\Delta x}{|v_{\max}|}$. The numerical scheme is based on a splitting method: Given an initial datum $f(x,v,0) = f_0(x,v)$, we split the system at every time $t_k = k\Delta t$ into the following steps:

21
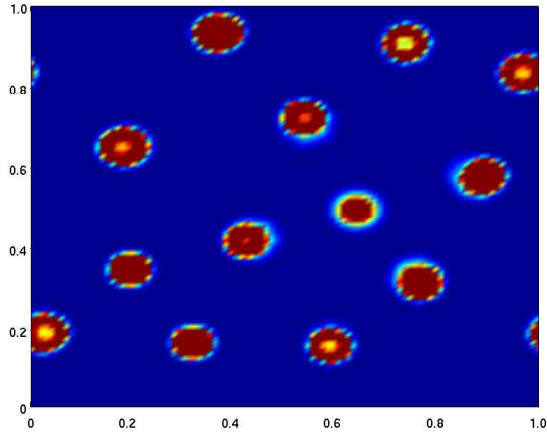
Figure 6: The first-order mean-field model in a periodic 2D setting with random initial condition (upper left panel). After the initial rapid smoothing of the high-frequency components, several ring-shaped structures are created (upper right panel), which eventually turn into an almost regular pattern of well localized aggregates (lower left panel). However, the smaller aggregates may be unstable and diffusively disintegrate, and their mass is absorbed by their neighbors, (lower right panel).

R=0.06   R=0.11
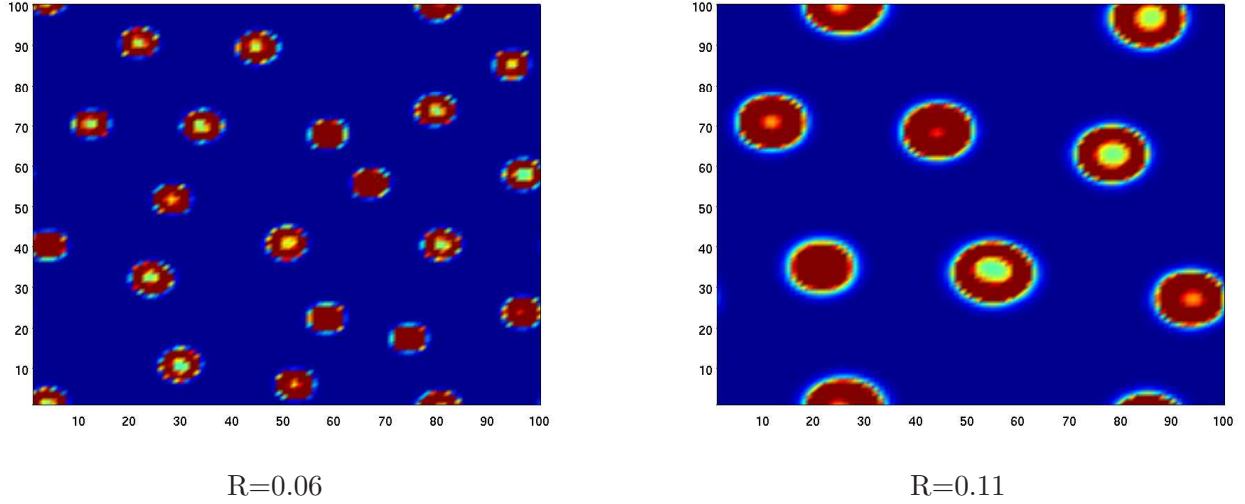
Figure 7: An example of patterns produced by the first-order mean-field model in a periodic 2D setting with the sampling radii $R = 0.6$ (left panel) and $R = 0.11$ (right panel).

1. Solve transport equation in x

$$\frac{\partial f}{\partial t} + v_i \cdot \nabla_x f = 0, \tag{5.1}$$

   subject to periodic boundary conditions on Ω, for every v_i on the time interval $t \in [t_k, t_k + \frac{1}{2}\Delta t]$, using an upwind scheme with the superbee flux limiter.

2. Starting with the solution of the transport equation (5.1), solve

$$\frac{\partial f}{\partial t} = \nabla_v \cdot \left(H(W \circledast f)vf + \frac{1}{2}\nabla_v(G(W \circledast f)^2 f) \right), \tag{5.2}$$

   with no flux boundary conditions on V, on the time interval $[t_k, t_{k+1}]$, using a semi-implicit time discretization.

3. Finally solve (5.1) using the solution of (5.2) for another half time step $\frac{1}{2}\Delta t$.

We set the computational domain to $\Omega = [0,1]$, the velocity domain to $V = [-1,1]$ and the mesh sizes to $\Delta x = 10^{-2}$ and $\Delta v = 2 \times 10^{-2}$. We choose similar conditions as in the individual based model, i.e. a limited cone of vision $W(x,v) = w(|x|, \frac{x \cdot v}{|x||v|})$ with $w(s,z) = \chi_{[0,R]}(s)\chi_{[0,1]}(z)$. The sampling radius is set to $R = 0.07$, $G(s) = \exp(-2s)$ and $H \equiv 2$. The initial datum corresponds to a small perturbation of a uniform distribution with mass one. Snapshots of the evolution of the particle distribution density $f(t,x,v)$ and the mass density $\rho(t,x)$ at different times are depicted in Fig. 8. We observe a fast smoothing of $f(t,x,v)$ in time and a subsequent formation of a stable aggregate.
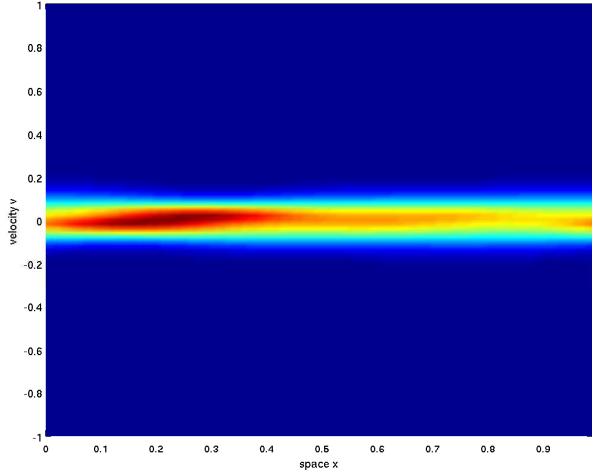
If we decrease the sampling radius to $R = 0.03$, two separate aggregates form, see Fig. 9. Again, we observe that with a smaller sampling radius the aggregation happens on a faster time scale.
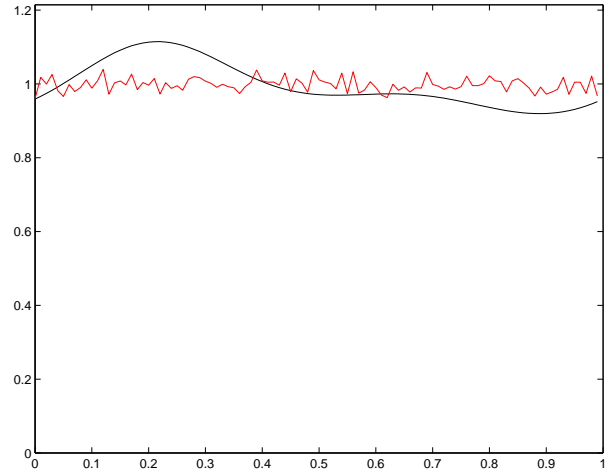
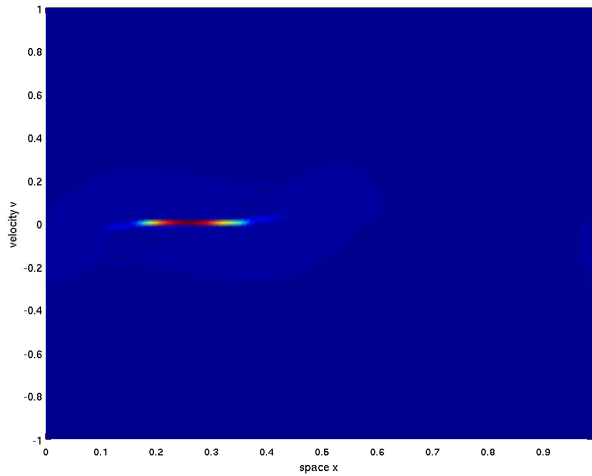(a) Particle distribution function $f(t, x, v)$ at $t = 4$
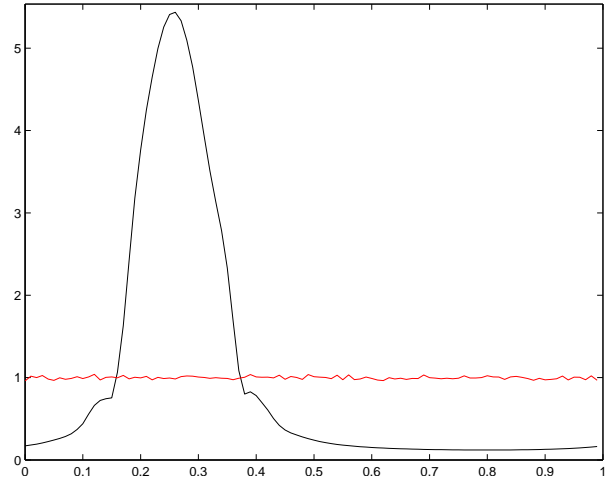

(b) Mass density $\rho(t, x)$ at $t = 4$

(c) Particle distribution function $f(t, x, v)$ at $t = 8$
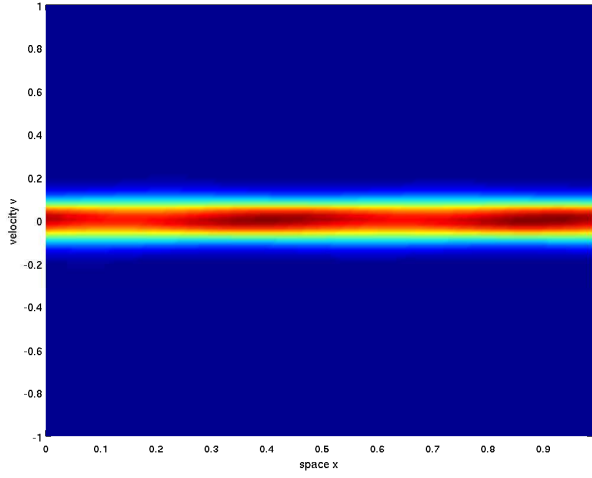

(d) Mass density $\rho(t, x)$ at $t = 8$
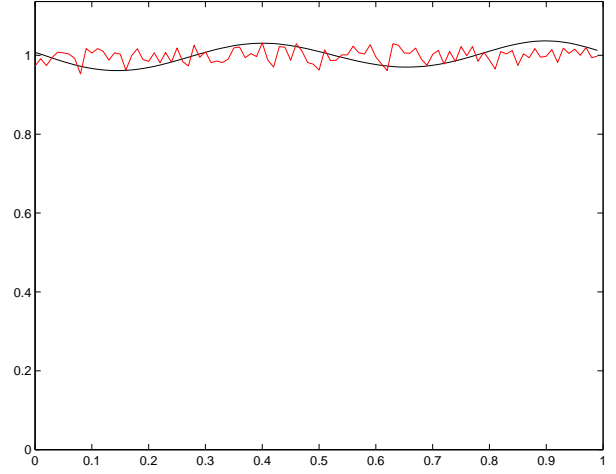

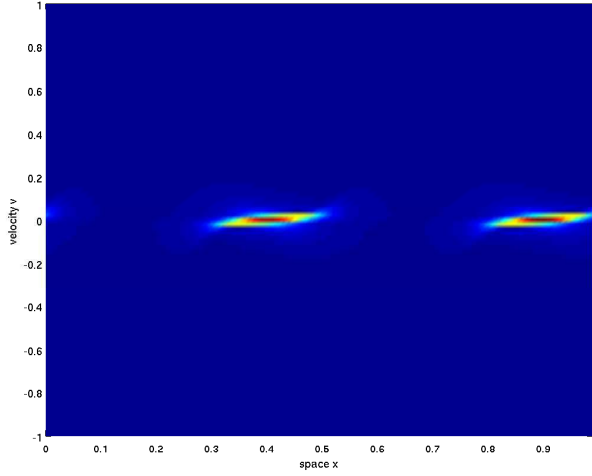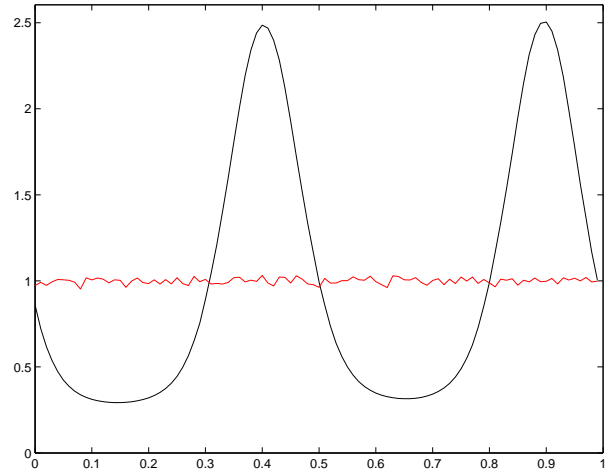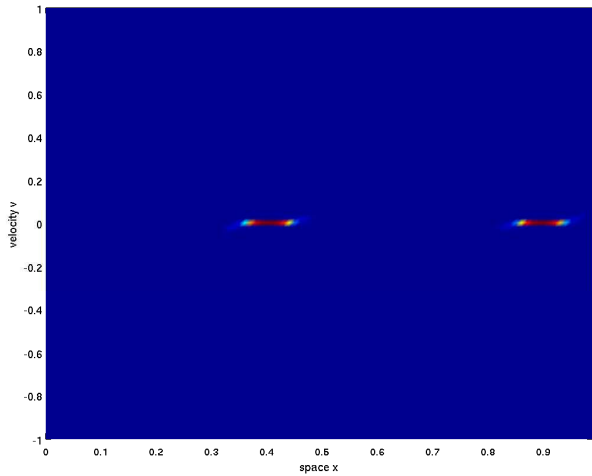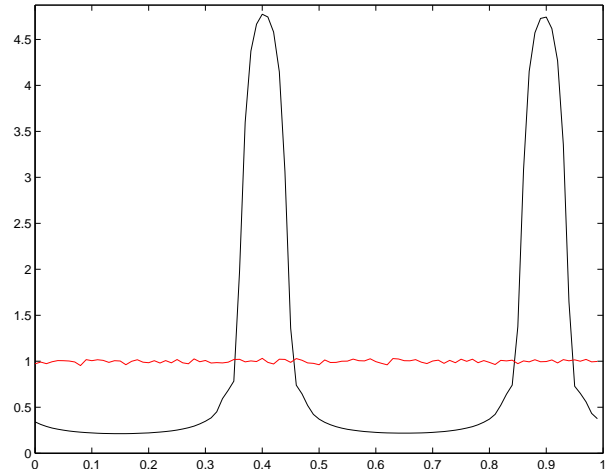(e) Particle distribution function $f(t, x, v)$ at $t = 20$

(f) Mass density $\rho(t, x)$ at $t = 20$

24

Figure 8: Second order model with limited vision and $R = 0.07$; the red line in the left column indicates the initial datum.

(a) Particle distribution function $f(t,x,v)$ at $t=4$

(b) Mass density $\rho(t,x)$ at $t=4$

(c) Particle distribution function $f(t,x,v)$ at $t=8$

(d) Mass density $\rho(t,x)$ at $t=8$

(e) Particle distribution function $f(t,x,v)$ at $t=20$

(f) Mass density $\rho(t,x)$ at $t=20$

Figure 9: Second order model with limited vision and $R = 0.03$; the red line in the left column indicates the initial datum.

References

[1] K. Anguige, C. Schmeiser: A one-dimensional model of cell diffusion and aggregation, incorporating volume filling and cell-to-cell adhesion. J. Math. Biology (2009), pp.

[2] B. Briscoe, M. Lewis and S. Parrish: Home range formation in wolves due to scent marking. Bulletin of Mathematical Biology 64 (2002) 64, 261–284.

[3] M. Burger: Nonstandard BBGKY hierarchies. Preprint, 2011.

[4] H. Brezis, *Analyse fonctionelle*, Dunod, Paris, 1999.

[5] L. C. Evans, Partial Differential Equations,1998, American Mathematical Society, Providence, Rhode Island.

[6] P. Grindrod: Models of individual aggregation in single and multispecies communities. J. Math. Biol. 26 (1988), pp. 651–660.

[7] D. Grünbaum and A. Okubo: Modelling social animal aggregations. In: S. A. Levin (Ed.), Frontiers of Theoretical Biology. Vol. 100 of Lecture Notes in Biomathematics (1994), Springer-Verlag.

[8] E. Hernández-García and C. López: Clustering, advection, and patterns in a model of population dynamics with neighborhood-dependent rates. Phys. Rev. E 70 (2004), pp. 016216.

[9] R. Jeanson, S. Blanco, R. Fournier, J.-L. Deneubourg, V. Fourcassié and G. Theraulaz: A model of animal movements in a bounded space, Journal of Theoretical Biology 225 (2003), pp. 443451.

[10] R. Jeanson, C. Rivault, J.-L. Deneubourg, S. Blanco, R. Fournier, C. Jost and G. Theraulaz: Self-organised aggregation in cockroaches, Animal Behaviour 69 (2005), pp. 169180.

[11] C. Jost, R. Jeanson, J. Gautrais, B.-R. Bengoudifa, G. Theraulaz: Sensitivity of cockroach aggregation to individual and external parameters. Preprint, 2011.

[12] K. Kawasaki: Diffusion and the formation of spatial distribution. Mathematical Sciences 16 (1978), 47–52.

[13] E. F. Keller and L. A. Segel: Initiation of slime mold aggregation viewed as an instability. J. Theoret. Biol. 26 (1970), pp. 399–415.

[14] J. Krause and G. Ruxton: Living in groups. Oxford University Press, 2002.

[15] J. Krebs and N. Davies (eds.): Behavioural Ecology: An Evolutionary Approach. Sinauer, Sunderland Massachussetts, 1984.

[16] C. López: Self-propelled nonlinearly diffusing particles: Aggregation and continuum description. Phys. Rev. E 72 (2005), pp. 061109.

[17] C. López: Macroscopic description of particle systems with nonlocal density-dependent diffusivity. Phys. Rev. E 74 (2006), pp. 012102.

[18] P. Maini, L. Malaguti, C. Marcelli and S. Matucci: Diffusion-aggregation processes with monostable reaction terms. Discrete and Continuous Dynamical Systems Series B, 6 (5), 2006, 1175–1189.

[19] M. Mimura and M. Yamaguti: Pattern formation in interacting and diffusing systems in population biology. Advances in Biophysics 15 (1982), pp. 19–65.

[20] A. Mogilner and L. Edelstein-Keshet: A non-local model for a swarm. J. Math. Biol. 38 (1999), 534–570.

[21] J. Murray: Mathematical Biology. Springer, 2001.

[22] A. Okubo: Diffusion and Ecological Problems: Mathematical Models. Springer Verlag, 1980.

[23] V. Padrón: Aggregation on a Nonlinear Parabolic Functional Diferential Equation. Divulgaciones Matemáticas v. 6, No. 2 (1998), 149–164.

[24] V. Padrón: Effect of aggregation on population recovery modeled by a forward-backward pseudoparabolic equation. Trans. Amer. Math. Soc. 356 (2004), 2739–2756.

[25] F. Sánchez-Garduño, P. Maini and J. Pérez-Velázquez: A non-linear degenerate equation for direct aggregation and traveling wave dynamics. Discrete and Continuous Dynamical Systems Series B, 13 (2), 2010, pp. 455–487.

[26] M. Rust, J. Owens and D. Reierson: Understanding and controlling the german cockroach. Oxford University Press, 1995.

[27] M. J. Schnitzer: Theory of continuum random walker and application to chemotaxis. Phys. Rev. E 48 (1993), pp. 2553–2568.

[28] J. Simon: Compact sets in the space $L^p(0, T; B)$. Ann. Mat. Pure Appl. IV (146), 1987, pp. 65-96.

[29] J. Skellam: Random dispersal in theoretical populations. Biometrika 38 (1951), 196218.

[30] C. Topaz, A. Bertozzi and M. Lewis: A nonlocal continuum model for biological aggregation.

[31] P. Turchin: Population Consequences of Aggregative Movement. J. of Animal Ecology, Vol. 58, No. 1 (1989), pp. 75–100.

[32] P. Turchin, P. Kareiva: Aggregation in *Aphis varians*: an effective strategy for reducing predation risk. Ecology 70 (1989), pp. 1008-1016.